

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日:

2004年1月29日(29.01.2004)

PCT

(10) 国际公布号:

WO 04/010702 A1

(51) 国际分类号⁷: H04N 7/26

(21) 国际申请号: PCT/CN02/00673

(22) 国际申请日: 2002年9月23日(23.09.2002)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
02125274.2 2002年7月22日(22.07.2002) CN(71) 申请人(对除美国以外的所有指定国): 中国科学院
计算技术研究所(INSTITUTE OF COMPUTING
TECHNOLOGY CHINESE ACADEMY OF
SCIENCES) [CN/CN]; 中国北京市海淀区中关村
科学院南路6号中科院计算所数字化室(物理所),
Beijing 100080 (CN)。

(72) 发明人: 及

(75) 发明人/申请人(仅对美国): 高文(GAO, Wen)
[CN/CN]; 马思伟(MA, Siwei) [CN/CN]; 吕岩(LU,
Yan) [CN/CN]; 中国北京市海淀区中关村科学院
南路6号中科院计算所数字化室(物理所), Beijing
100080 (CN)。

(74) 代理人:

北京北新智诚知识产权代理有限公司(BEIJING

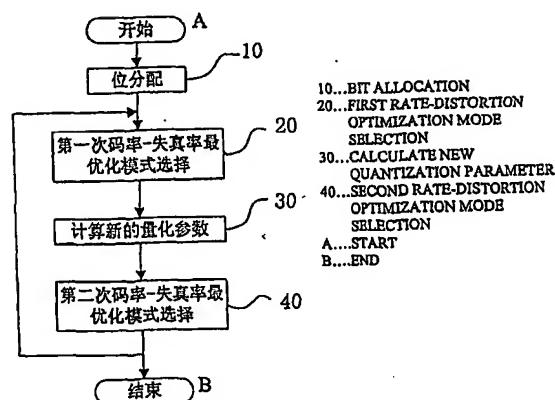
BEIXIN-ZHICHENG INTELLECTUAL
PROPERTY AGENT CO., LTD.); 中国北京市
西城区西直门内南大街16号, Beijing 100035 (CN)(81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA,
BB, BG, BR, BY, BZ, CA, CH, CO, CR, CU, CZ, DE,
DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,
LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO,
RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT,
TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW(84) 指定国(地区): ARIPO专利(GH, GM, KE, LS, MW,
MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM,
AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT,
BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,
GR, IE, IT, LU, MC, NL, PT, SE, SK, TR),
OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG)

本国际公布:

— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期
PCT公报期刊起始的“代码及缩写符号简要说明”。(54) Title: A BIT-RATE CONTROL METHOD AND DEVICE COMBINED WITH RATE-DISTORTION
OPTIMIZATION

(54) 发明名称: 组合率失真优化的码率控制的方法及其装置



(57) Abstract: A bit-rate control method combined with rate-distortion optimization, first, allocating corresponding bit numbers for each frame in a GOP; then, obtaining quantization parameters for bit rate control by performing rate-distortion optimization mode selection for each macro block in a frame; a bit-rate control device combined with rate-distortion optimization, comprises a JVT processing module, a rate-distortion optimization mode selection module, a virtual buffer module, and a whole complexity estimate module; by said method and device, the disadvantage of conventional bit-rate control method in the RDO based video encoding is overcome, because of having synthetically considered both RDO and bit-rate control, a bit-rate control scheme is provided for RDO based video encoding technology, therefore while controlling the bit-rate, it ensures the visual quality of decoded image. In addition, the present invention provides an embodiment which utilizes JVT standard being established, and provides an embodied technical path for resolving the bit-rate control problem for JVT standard.

[见续页]



(57) 摘要

一种结合率失真优化的码率控制的方法，首先，为 GOP 中各帧分配相应的比特数；然后，针对帧内各宏块进行率失真最优化模式选择，获得码率控制的量化参数；一种结合率失真优化的码率控制的装置，包括 JVT 处理模块、率失真最优化模式选择模块、虚拟缓冲区模块和全局复杂性估计模块；通过上述方法及其装置，克服了传统的码率控制方法在基于 RDO 的视频编码中的缺陷，由于综合考虑了 RDO 与码率控制，为基于 RDO 的视频编码技术提供了码率控制方案；因此在控制码率的同时，保证了解码图像的视觉质量。另外，本发明提供了一种采用正在制定的 JVT 标准作为实施的案例，为 JVT 标准提供了解决码率控制问题的具体技术路径。

结合率失真优化的码率控制的方法及其装置

技术领域

本发明涉及一种结合率失真优化的码率控制的方法及其装置，特别涉及一种适于流媒体传输的视频编解码技术和装置，属于数字视音频编码标准化技术和数字媒体处理技术领域。

背景技术

高效的视频编解码技术是实现多媒体数据存储与传输的关键，而先进的视频编解码技术通常以标准的形式存在。目前，典型的视频压缩标准有：国际标准化组织（ISO）下设的运动图像专家组（Moving Picture Expert Group，简称 MPEG）推出的 MPEG 系列国际标准、国际电信联盟（ITU）提出的 H.26x 系列视频压缩标准、以及 ISO 和 ITU 建立的联合视频工作组（Joint Video Team，简称 JVT）正在制定的 JVT 视频编码标准等。

JVT 标准采用的是一种新型的编码技术，它比现存的任何一种编码标准的压缩效率都要高的多。JVT 标准在 ISO 中的正式名称是 MPEG-4 标准的第十部分，在 ITU 中的正式名称是 H.264 标准。

码率控制是实现恒定码率的视频编码的关键技术，如果没有码率控制，任何视频编码标准的应用都几乎是不可能的。举例来说：在信道传输过程中，如果没有码率控制就只能得到码率不均匀的流媒体，进而产生两方面的问题：一方面，当码率超过信道的传输能力时会造成传输阻塞，从而引发“数字悬崖”；另一方面，当码率低于信道传输速率时会造成信道带宽的浪费，进而影响重构视频的质量。

基于上述的原因，任何视频编码标准在制定过程中都会给出一套适用于其本身的码率控制方案，例如：MPEG-2 标准中的测试模型 5（Test Model 5，简称 TM5）、H.263 标准中的测试模型 8（Test Model Near 8，简称 TMN8）等。

结合率失真优化（Rate-Distortion Optimization，以下简称 RDO）技术的视频编码是目前比较先进的方法。例如，正在制定的 JVT 编码标准中就引入了 RDO 技术。RDO 技术为编码帧的每一宏块选择最优的运动向量与编码模式，使得码率与失真率达到最好的折衷关系。

RDO 为每一宏块选择运动向量与编码模式时受到该宏块的量化参数的影响，而宏块级的码率控制就是为每一宏块选择一个合适的量化参数；或者说，为码率控制选择量化参数需要得到通过 RDO 获得编码模式，而 RDO 在选择模式时需要预先知道量化参数。因此，对于采用 RDO 技术的视频编码方法，传统的码率控制技术已经不再有效。这也是目前 JVT 标准还没有给出码率控制技术的原因。此外，采用 RDO 技术使得帧内编码帧（Intra Frame，简称 I 帧）、预测编码帧（Prediction Frame，简称 P 帧）和双向预测帧（Bidirectional Prediction Frame，简称 B 帧）的帧间复杂性比例发生变化，因此传统的位分配模型以及适应性量化机制也必须进行改进。

发明内容

本发明的主要目的在于提供一种结合率失真优化的码率控制的方法及其装置, 其克服传统的码率控制方法在基于 RDO 的视频编码中的缺陷, 综合考虑 RDO 与码率控制, 为基于 RDO 的视频编码技术提供码率控制方案, 在控制码率的同时, 保证解码图像的视觉质量。

本发明的另一目的在于提供一种结合率失真优化的码率控制的方法及其装置, 采用正在制定的 JVT 标准作为实施案例, 为 JVT 标准提供解决码率控制问题的具体技术路径。

本发明的目的是这样实现的:

一种结合率失真优化的码率控制的方法, 至少包括如下的步骤:

步骤 10: 为图片组 (Group of Picture, 简称 GOP) 中各帧分配相应的比特数;

步骤 20: 针对帧内各宏块进行率失真最优化模式选择, 获得码率控制的量化参数。

上述步骤 10 的具体操作为: 在对 GOP 中各帧进行编码之前, 根据平均帧的大小为 GOP 中各帧分配相应的比特数。

所述的平均帧的大小根据如下公式计算:

$R/F = R \div F$, 其中:

R 为码率控制的目标码率;

F 为编码帧率;

R/F 为平均帧的大小的值。

所述的为 GOP 中各帧分配相应的比特数还进一步包括: 在该 GOP 中的各帧之间进行协调, 具体的协调处理依据如下的公式:

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{\text{bit_rate}}{8 \times \text{picture_rate}} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_p X_b}{K_b X_p}}, \frac{\text{bit_rate}}{8 \times \text{picture_rate}} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_b X_p}{K_p X_b}}, \frac{\text{bit_rate}}{8 \times \text{picture_rate}} \right\}$$

其中,

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数;

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目;

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值, 其为编码一帧所用的比特数与平均量化参数的乘积;

bit_rate 为目标码率;

$picture_rate$ 为编码帧率;

K_p 和 K_b 为经验值, 分别为 P, B 帧与 I 帧间的复杂性比例, 其取值范围为不小于 1.0 且不大于 2.0;

R 为当前 GOP 剩余的比特数; 且在编码一帧后, 依据如下公式对 R 进行更新:

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

在对一个 GOP 进行编码之前, 还依据如下公式对为当前 GOP 剩余的比特数进行初始化;

$$R = G + R_{prev}$$

$$G = bit_rate \times N \div picture_rate$$

其中:

R 为当前 GOP 剩余的比特数,

N 为当前 GOP 所包含的帧数,

G 为码率= bit_rate 时, 一个包含 N 帧图像的 GOP 的比特数,

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数, 且对于第一个 GOP,

$$R_{prev} = 0;$$

bit_rate 为目标码率;

$picture_rate$ 为编码帧率。

上述的 X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定:

$$X_i = a \times bit_rate$$

$$X_p = b \times bit_rate$$

$$X_b = c \times bit_rate$$

其中:

a 、 b 、 c 为经验值, 其取值范围分别为: 100/115-160/115、20/115-60/115、10/115-42/115;

bit_rate 为目标码率。

所述的步骤 20 至少包括第一次率失真最优化模式选择, 即: 用当前宏块之前一宏块的量化参数作为预测值进行模式选择; 具体为: 依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

其中:

s 为源图像当前宏块的辉度值;

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{\frac{v_m-1}{3}}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{\frac{v_m-1}{3}}$

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

当对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m-p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和;

$R(m-p)$ 为编码运动向量所用的比特数;

s 为源图像当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

当得到上述的码率控制的量化参数后, 即第一次率失真最优化模式选择之后, 该码率控制的量化参数即可被输出而直接用于实际的码率控制之中。

为了进一步优化码率控制的量化参数, 在第一次率失真最优化模式选择之后还进一步包括: 步骤 30: 当前宏块量化参数的确定; 具体为: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

根据当前宏块的活动性调整宏块的量化参数的处理具体为: 采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或运动补偿后的绝对值和作为宏块的活动性标志, 当前宏块的空间活动性依据下式计算:

$$N_{act_m} = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中,

i 为像素的横坐标, 并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标;

j 为像素的纵坐标, 并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标;

N_act_m 为当前宏块的空间活动性;

$s(i, j)$ 为源像素 (i, j) 的辉度值;

$c(i, j)$ 为像素 (i, j) 的预测值;

avg_act 为先前编码的同类型帧的平均 act_m 活动性;

act_m 为经过运动补偿或帧内预测后的当前宏块的绝对值和。

所述的建立虚拟缓冲区模型具体为: 建立虚拟缓冲区的饱和度与宏块量化参数的映射关系; 其中, 当前宏块的量化参数由如下公式确定:

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_act_m$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times bit_rate / picture_rate$$

其中,

Q_m 为当前宏块的量化参数值;

d_m^n 为编码当前宏块时当前虚拟缓冲区的饱和度; 其中, 对于 I, P, B 帧, 其取值分别为 d_m^i 、 d_m^p 、 d_m^b ;

B_{m-1} 为编码前一宏块所用的比特数;

d_0^n 为编码该帧时虚拟缓冲区的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b

r 为虚拟缓冲区的大小值;

虚拟缓冲区最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m=MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

所述的虚拟缓冲区饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区的初始饱和度;

K_p 为 I, P 帧复杂性比例常数;

K_b 为 I, B 帧复杂性比例常数。

在确定当前宏块的量化参数后, 还进一步包括: 步骤 40: 第二次率失真最优化模式选择, 即: 用确定后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择; 具体依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{\frac{QP-1}{3}}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{\frac{QP-1}{3}}$

$D(s, c, MODE | QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE | QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

当对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m - p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和。 $R(m-p)$ 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中, λ_{MODE} 为拉格朗日常数。

经过上述的处理, 得到的码率控制的量化参数一方面可以输出, 用于实际的码率控制; 另一方面, 还可以作为再进一步的率失真最优化模式选择的初始参数, 以获得更为精确的输出。

一种结合率失真优化的码率控制的装置, 至少包括 JVT 处理模块、率失真最优化模式选择模块、虚拟缓冲区模块和全局复杂性估计模块; 其中,

JVT 处理模块输入帧数据, 并与率失真最优化模式选择模块、虚拟缓冲区模块和全局复杂性估计模块连接;

该率失真最优化模式选择模块分别与虚拟缓冲区和全局复杂性估计模块连接, 它从 JVT 处理模块输入信号, 并根据虚拟缓冲区模块和全局复杂性估计模块输出的信号将其输出信号回送给 JVT 处理模块, 并经过 JVT 处理模块处理后输出相应的量化参数信号。

所述的位分配为: 帧数据输入时, 首先为 GOP 中各帧分配相应的的比特数; 具体是:

在对 GOP 中各帧进行编码之前, 根据平均帧的大小为 GOP 中各帧分配相应的比特数。

所述的平均帧的大小根据如下公式计算:

$R/F = R \div F$, 其中:

R 为码率控制的目标码率;

F 为编码帧率;

R/F 为平均帧的大小的值。

所述的为 GOP 中各帧分配相应的比特数还包括: 在该 GOP 中的各帧之间进行协调, 具体的协调处理依据如下的公式:

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_b X_p}{K_p X_b}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_p X_b}{K_b X_p}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

其中,

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数;

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目;

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值, 其为编码一帧所用的比特数与平均量化参数的乘积;

bit_rate 为目标码率;

$picture_rate$ 为编码帧率;

K_p 和 K_b 为经验值, 分别为 P、B 帧与 I 帧间的复杂性比例, 其取值范围为不小于 1.0 且不大于 2.0;

R 为当前 GOP 剩余的比特数; 且在编码一帧后, 依据如下公式对 R 进行更新:

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

在对一个 GOP 进行编码之前, 还依据如下公式对为当前 GOP 剩余的比特数进行初始化;

$$R = G + R_{prev}$$

$$G = \text{bit_rate} \times N \div \text{picture_rate}$$

其中:

R 为当前 GOP 剩余的比特数,

N 为当前 GOP 所包含的帧数,

G 为码率= bit_rate 时, 一个包含 N 帧图像的 GOP 的比特数,

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数, 且对于第一个 GOP,

$$R_{\text{prev}} = 0;$$

bit_rate 为目标码率;

picture_rate 为编码帧率。

上述的 X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定:

$$X_i = a \times \text{bit_rate}$$

$$X_p = b \times \text{bit_rate}$$

$$X_b = c \times \text{bit_rate}$$

其中:

a 、 b 、 c 为经验值, 其取值范围分别为: 100/115-160/115、20/115-60/115、10/115-42/115;

bit_rate 为目标码率。

率失真最优化模式选择模块采用当前宏块之前一宏块的量化参数作为预测值进行模式选择; 其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

其中:

s 为源图像当前宏块的辉度值;

$$D(s, c, \text{MODE} | QP) + \lambda_{\text{MODE}} R(s, c, \text{MODE} | QP)$$

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{\text{MODE}} = 0.85 \times 2^{\frac{QP-1}{3}}$

对于 B 帧, $\lambda_{\text{MODE}} = 4 \times 0.85 \times 2^{\frac{QP-1}{3}}$

$D(s, c, \text{MODE} | QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, \text{MODE} | QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

率失真最优化模式选择模块对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{\text{MOTION}}) = SA(T) D(s, c(m)) + \lambda_{\text{MOTION}} R(m - p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

SA(T)D 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和。R(m-p) 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值,

c 为预测图像的辉度值,

m 为运动向量,

p 为运动向量的预测值。

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

当得到上述的码率控制的量化参数后, 即第一次率失真最优化模式选择之后, 该码率控制的量化参数即可被输出而直接用于实际的码率控制之中。

为了进一步优化码率控制的量化参数, 在率失真最优化模式选择模块第一次率失真最优化模式选择之后, 其输出给 JVT 处理模块的信号还由 JVT 处理模块进一步当前宏块的量化参数进行确定; 具体为: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

上述的根据当前宏块的活动性调整宏块的量化参数的处理具体为: 率失真最优化模式选择模块采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或运动补偿后的绝对值和作为宏块的活动性标志,

当前宏块的空间活动性依据下式计算:

$$N_{act_m} = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中,

i 为像素的横坐标, 并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标;

j 为像素的纵坐标, 并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标;

N_{act_m} 为当前宏块的空间活动性;

$s(i, j)$ 为源像素 (i, j) 的辉度值;

$c(i, j)$ 为像素 (i, j) 的预测值;

avg-act 为先前编码的同类型帧的平均 act_m 活动性;

act_m 为经过运动补偿或帧内预测后当前宏块的绝对值和。

所述的建立虚拟缓冲区模型具体为: 模块 2 建立虚拟缓冲区的饱和度与宏块量化参数的映射关系; 其中, 当前宏块的量化参数由如下公式确定:

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_{act_m}$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times bit_rate / picture_rate$$

其中,

Q_m 为当前宏块的量化参数值;

d_m^n 为编码当前宏块时当前虚拟缓冲区的饱和度; 其中, 对于 I, P, B 帧, 其取值分别为 d_m^i 、 d_m^p 、 d_m^b ;

B_{m-1} 为编码前一宏块所用的比特数;

d_0^n 为编码该帧时虚拟缓冲区的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b

r 为虚拟缓冲区的大小值;

虚拟缓冲区最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m = MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

所述的虚拟缓冲区饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区的初始饱和度;

K_p 为 I, P 帧复杂性比例常数;

K_b 为 I, B 帧复杂性比例常数。

在确定当前宏块的量化参数后, JVT 处理模块将该参数信号输出给率失真最优化模式选择模块, 由该率失真最优化模式选择模块进行率失真最优化模式选择, 即: 用确定后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择; 其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{u_{m-1}/3}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{u_{m-1}/3}$

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

当率失真最优化模式选择模块对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m-p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和。 $R(m-p)$ 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

率失真最优化模式选择模块根据虚拟缓冲区和全局复杂性估计模块输出的信号, 将处理后的量化参数信号回送给 JVT 处理模块, 并经过 JVT 处理模块处理后输出。

通过上述方法及其装置, 克服了传统的码率控制方法在基于 RDO 的视频编码中的缺陷, 由于综合考虑了 RDO 与码率控制, 为基于 RDO 的视频编码技术提供了码率控制方案, 因此在控制码率的同时, 保证了解码图像的视觉质量。另外, 本发明提供了一种采用正在制定的 JVT 标准作为实施的案例, 为 JVT 标准提供了解决码率控制问题的具体技术路径。

附图说明

图 1 为本发明一实施例的流程图。

图 2 为本发明装置的原理结构图。

具体实施方式

以下结合附图和具体的实施例对本发明作进一步的详细说明:

实施例 1:

本发明的一个具体的实施方案是: 假定码率控制的目标码率为 R , 编码帧率为 F , 则平均每帧的大小为 R/F 。在编码一帧图像之前, 根据 R/F 值为该帧分配一定的比特数; 对复杂性不同的 I 帧、P 帧和 B 帧分配不同的比特数, 以保证解码图像的质量。

然后, 对帧内的每一宏块进行一次 RDO 最优模式选择。其中, 量化参数的选择以前一宏块做为参考。经过本次 RDO 以后, 可以得到该宏块经过帧内预测或运动补偿后的绝对值

和 (Sum of Absolute Difference, 简称 SAD), 并用这个 SAD 值标志宏块的活动性。

接下来, 根据虚拟缓冲区的饱和度以及当前宏块的活动性, 计算当前宏块的量化参数。

最后, 采用上述的量化参数进行第二次 RDO 选择最优的编码模式。

具体参见附图 1, 本发明提出的码率控制方法的一个具体实施例包含如下步骤:

步骤 10: 位分配。首先为一个图片组 (Group of Picture, 简称 GOP) 分配一定的比特数, 然后在这个 GOP 中的各帧之间进行协调。

本步骤的具体处理过程包括:

在对 GOP 中各帧进行编码之前, 根据平均帧的大小为 GOP 中各帧分配相应的的比特数。

平均帧的大小根据如下公式计算: $R/F=R \div F$; 其中:

R 为码率控制的目标码率;

F 为编码帧率;

R/F 为平均帧的大小的值。

所述的为 GOP 中各帧分配相应的的比特数还进一步包括: 在该 GOP 中的各帧之间进行协调, 具体的协调处理依据如下的公式:

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_p X_p}{K_p X_b}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_b X_b}{K_b X_p}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

其中,

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数;

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目;

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值, 其为编码一帧所用的比特数与平均量化参数的乘积;

bit_rate 为目标码率;

$picture_rate$ 为编码帧率;

K_p 和 K_b 为经验值, 分别为 P, B 帧与 I 帧间的复杂性比例, 其取值范围为不小于 1.0 且

不大于 2.0;

R 为当前 GOP 剩余的比特数; 且在编码一帧后, 依据如下公式对 R 进行更新:

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

在对一个 GOP 进行编码之前, 还依据如下公式对为当前 GOP 剩余的比特数进行初始化;

$$R = G + R_{prev}$$

$$G = bit_rate \times N \div picture_rate$$

其中:

R 为当前 GOP 剩余的比特数,

N 为当前 GOP 所包含的帧数,

G 为码率= bit_rate 时, 一个包含 N 帧图像的 GOP 的比特数,

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数, 且对于第一个 GOP,

$$R_{prev} = 0;$$

bit_rate 为目标码率;

$picture_rate$ 为编码帧率。

上述的 X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定:

$$X_i = a \times bit_rate$$

$$X_p = b \times bit_rate$$

$$X_b = c \times bit_rate$$

其中:

a 、 b 、 c 为经验值, 其取值范围分别为: 100/115-160/115、20/115-60/115、10/115-42/115;

bit_rate 为目标码率。

步骤 20: 进行第一次率失真最优化计算。即: 在为当前宏块进行第一次基于 RDO 的模式选择时, 先用前一宏块的量化参数作为预测值进行模式选择。这是因为: 宏块间的量化参数波动会影响图像的质量, 尤其是在低码率这种波动的影响尤为严重。

本步骤的具体处理包括:

用当前宏块之前一宏块的量化参数作为预测值进行模式选择; 即: 依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

其中:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{v_{m-1}/3}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{v_{m-1}/3}$

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

当上述的处理为对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m-p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和。 $R(m-p)$ 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值,

c 为预测图像的辉度值,

m 为运动向量,

p 为运动向量的预测值。

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

步骤 30: 计算当前宏块的量化参数。首先计算当前宏块的活动性, 然后根据虚拟缓冲区模型建立了虚拟缓冲区的饱和度与宏块量化参数的映射关系。其中, 宏块的量化参数由当前虚拟缓冲区的饱和度以及宏块的活动性决定。

本步骤的具体处理为:

在上述的第一次率失真最优化模式选择之后, 还进一步对当前宏块量化参数进行确定; 即: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

首先, 根据当前宏块的活动性调整宏块的量化参数的处理具体为: 采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或运动补偿后的绝对值和作为宏块的活动性标志, 当前宏块的空间活动性依据下式计算:

$$N_{act_m} = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中,

i 为像素的横坐标, 并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标;

j 为像素的纵坐标, 并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标;

N_act_m 为当前宏块的空间活动性;

$s(i, j)$ 为源像素 (i, j) 的辉度值;

$c(i, j)$ 为像素 (i, j) 的预测值;

avg_act 为先前编码的同类型帧的平均 act_m 活动性;

act_m 为经过运动补偿或帧内预测后当前宏块的绝对值和。

其次, 建立虚拟缓冲区模型具体为: 建立虚拟缓冲区的饱和度与宏块量化参数的映射关系; 其中, 当前宏块的量化参数由如下公式确定:

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_act_m$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times bit_rate / picture_rate$$

其中,

Q_m 为当前宏块的量化参数值;

d_m^n 为编码当前宏块时当前虚拟缓冲区的饱和度; 其中, 对于 I, P, B 帧, 其取值分别为 d_m^i 、 d_m^p 、 d_m^b ;

B_{m-1} 为编码前一宏块所用的比特数;

d_0^n 为编码该帧时虚拟缓冲区的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b

r 为虚拟缓冲区的大小值;

虚拟缓冲区最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m=MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

上述的虚拟缓冲区饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区的初始饱和度;

K_p 为 I, P 帧复杂性比例常数;

K_b 为 I, B 帧复杂性比例常数。

步骤 40: 进行第二次率失真最优化计算。采用步骤 30 计算得到的量化参数, 并采用步骤 20 的方法, 对当前宏块再次进行码率—失真率最优化模式选择, 找到当前宏块的最优编码模式。然后对该宏块进行量化编码, 并更新相关参数。

本步骤的具体处理为:

用确定后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择; 具体依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{Q_{-1}/3}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{Q_{-1}/3}$

$D(s, c, MODE | QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE | QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

当对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T) D(s, c(m)) + \lambda_{MOTION} R(m - p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T) D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和。

$R(m-p)$ 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

实施例 2:

参见附图 2, 针对 JVT 编码标准, 本实施例给出了实现上述方法的码率控制装置的原理结构框图。其输入 TSI 为原始的视频数据, 输出 TSO 为编码后的码流, 其中虚线框内为 JVT 源编码模块。包括以下主要模块: JVT 处理模块 4、率失真最优化模式选择模块 1、虚拟缓

冲区模块 2 和全局复杂性估计模块 3; 其中,

JVT 处理模块 4 输入帧数据, 并与率失真最优化模式选择模块 1、虚拟缓冲区模块 2 和全局复杂性估计模块 3 连接;

该率失真最优化模式选择模块 1 分别与虚拟缓冲区模块 2 和全局复杂性估计模块 3 连接, 它从 JVT 处理模块 4 输入信号, 并根据虚拟缓冲区模块 2 和全局复杂性估计模块 3 输出的信号将其输出信号回送给 JVT 处理模块 4, 并经过 JVT 处理模块 4 处理后输出相应的量化参数信号。

上述的位分配为: 帧数据输入时, 首先为 GOP 中各帧分配相应的的比特数; 具体是: 在对 GOP 中各帧进行编码之前, 根据平均帧的大小为 GOP 中各帧分配相应的的比特数。

所述的平均帧的大小根据如下公式计算:

$R/F = R \div F$, 其中:

R 为码率控制的目标码率;

F 为编码帧率;

R/F 为平均帧的大小的值。

上述的为 GOP 中各帧分配相应的的比特数还进一步包括: 在该 GOP 中的各帧之间进行协调, 具体的协调处理依据如下的公式:

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_p X_b}{K_b X_p}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_b X_p}{K_p X_b}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

其中,

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数;

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目;

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值, 其为编码一帧所用的比特数与平均量化参数的乘积;

bit_rate 为目标码率;

$picture_rate$ 为编码帧率;

K_p 和 K_b 为经验值, 分别为 P, B 帧与 I 帧间的复杂性比例, 其取值范围为不小于 1.0 且不大于 2.0;

R 为当前 GOP 剩余的比特数; 且在编码一帧后, 依据如下公式对 R 进行更新:

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

在对一个 GOP 进行编码之前, 还依据如下公式对为当前 GOP 剩余的比特数进行初始化;

$$R = G + R_{prev}$$

$$G = bit_rate \times N \div picture_rate$$

其中:

R 为当前 GOP 剩余的比特数,

N 为当前 GOP 所包含的帧数,

G 为码率= bit_rate 时, 一个包含 N 帧图像的 GOP 的比特数,

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数, 且对于第一个 GOP,

$$R_{prev} = 0;$$

bit_rate 为目标码率;

$picture_rate$ 为编码帧率。

上述的 X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定:

$$X_i = a \times bit_rate$$

$$X_p = b \times bit_rate$$

$$X_b = c \times bit_rate$$

其中:

a 、 b 、 c 为经验值, 其取值范围分别为: 100/115-160/115、20/115-60/115、10/115-42/115;

bit_rate 为目标码率。

上述的率失真最优化模式选择模块 1 采用当前宏块之前一宏块的量化参数作为预测值进行模式选择; 其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{u_{m-1}/3}$

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{u_{m-1}/3}$

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

该率失真最优化模式选择模块 1 对 P 帧或 B 帧进行估计时, 先经过 JVT 处理模块 4 的运动估计模块 SD 进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m - p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数,

$SA(T)D$ 为经过 JVT 处理模块 4 的运动补偿 SA (或并经过哈德曼变换) 后的宏块系数的绝对值和。 $R(m-p)$ 为编码运动向量所用的比特数,

s 为源图像当前宏块的辉度值,

c 为预测图像的辉度值,

m 为运动向量,

p 为运动向量的预测值。

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

在率失真最优化模式选择模块 1 第一次率失真最优化模式选择之后, 其输出给 JVT 处理模块 4 的信号还由 JVT 处理模块 4 进一步当前宏块的量化参数进行确定; 具体为: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

上述的根据当前宏块的活动性调整宏块的量化参数的处理具体为: 率失真最优化模式选择模块 1 采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或经过 JVT 处理模块 4 的运动补偿 SA 后的绝对值和作为宏块的活动性标志,

当前宏块的空间活动性依据下式计算:

$$N_{act_m} = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中,

i 为像素的横坐标, 并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标;

j 为像素的纵坐标, 并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标;

N_{act_m} 为当前宏块的空间活动性;

$s(i, j)$ 为源像素 (i, j) 的辉度值;

$c(i, j)$ 为像素 (i, j) 的预测值;

avg_act 为先前编码的同类型帧的平均 act_m 活动性;

act_m 为经过 JVT 处理模块 4 的运动补偿 SA 或帧内预测后的宏块 m 的绝对值和。

上述的建立虚拟缓冲区模型具体为: 由虚拟缓冲区模块 2 建立虚拟缓冲区的饱和度与宏块量化参数的映射关系; 其中, 当前宏块的量化参数由如下公式确定:

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_{act_m}$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times bit_rate / picture_rate$$

其中,

Q_m 为当前宏块的量化参数值;

d_m^n 为编码当前宏块时当前虚拟缓冲区模块 2 的饱和度; 其中, 对于 I, P, B 帧, 其取值分别为 d_m^i 、 d_m^p 、 d_m^b ;

B_{m-1} 为编码前一宏块所用的比特数;

d_0^n 为编码该帧时虚拟缓冲区模块 2 的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b

r 为虚拟缓冲区模块 2 的大小值;

虚拟缓冲区模块 2 最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m=MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

虚拟缓冲区模块 2 饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区模块 2 的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区模块 2 的初始饱和度;

K_p 为 I, P 帧复杂性比例常数;

K_b 为 I, B 帧复杂性比例常数。

在确定当前宏块的量化参数后, JVT 处理模块 4 将该参数信号输出给率失真最优化模式选择模块 1, 由该率失真最优化模式选择模块 1 进行率失真最优化模式选择, 即: 用确定

后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择；其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式：

$$D(s, c, \text{MODE} | QP) + \lambda_{\text{MODE}} R(s, c, \text{MODE} | QP)$$

其中：

s 为源图像当前宏块的辉度值；

c 为重构图像当前宏块的辉度值；

λ_{MODE} 为拉格朗日常数。

对于 I 帧或 P 帧， $\lambda_{\text{MODE}} = 0.85 \times 2^{\frac{QP-1}{3}}$

对于 B 帧， $\lambda_{\text{MODE}} = 4 \times 0.85 \times 2^{\frac{QP-1}{3}}$

$D(s, c, \text{MODE} | QP)$ 为衡量当前宏块压缩后的失真率；

$R(s, c, \text{MODE} | QP)$ 为编码当前宏块所用得比特数；

QP 为当前宏块的量化参数。

当率失真最优化模式选择模块 1 对 P 帧或 B 帧进行估计时，先经过 JVT 处理模块 4 的运动估计模块 SD 进行运动估计，并选择使下式最小的运动向量为当前宏块的运动向量：

$$J(m, \lambda_{\text{MOTION}}) = SA(T)D(s, c(m)) + \lambda_{\text{MOTION}} R(m-p)$$

其中，

$D(s, c(m))$ 为衡量运动补偿的失真率的函数，

$SA(T)D$ 为经过 JVT 处理模块 4 的运动补偿 SA （或并经过哈德曼变换）后的宏块系数的绝对值和。

$R(m-p)$ 为编码运动向量所用的比特数，

s 为源图像当前宏块的辉度值；

c 为预测图像的辉度值；

m 为运动向量；

p 为运动向量的预测值；

λ_{MOTION} 为拉格朗日常数，并且 $\lambda_{\text{MOTION}} = \sqrt{\lambda_{\text{MODE}}}$ ；其中，

λ_{MODE} 为拉格朗日常数。

最后，率失真最优化模式选择模块 1 根据虚拟缓冲区模块 2 和全局复杂性估计模块 3 输出的信号，将处理后的量化参数信号回送给 JVT 处理模块 4，并经过 JVT 处理模块 4 处理后输出。

综合上面所述，率失真最优化模式选择模块 1 用于计算当前宏块的量化参数，它的输入为虚拟缓冲区饱和度、宏块的空间活动性及帧的全局复杂性估计，输出为当前宏块的量化参数。

虚拟缓冲区模块 2 用于计算当前虚拟缓冲区的饱和度，它的输入为前一宏块编码产生的比特数，输出为编码当前宏块时虚拟缓冲区的饱和度。

全局复杂性模块 3 用于计算已编码帧的全局复杂性估计，它的输入为编码帧所用的比特数及编码帧平均的量化参数，输出为编码帧的全局复杂性估计。

再参见图 2，结合虚线框内 JVT 处理模块 4，针对一帧输入，本发明的装置对帧内的每一宏块进行处理。首先为当前宏块选择最优的编码模式，若宏块编码模式为帧内编码则开关 S 接 S1，加法器 a1 处进行帧内预测，然后对预测后的宏块系数进行离散余弦变换及量化，再对量化后的系数经过 JVT 处理模块 4 的熵编码模块 JN 进行熵编码。同时，JVT 处理模块 4 对这些系数进行逆量化及反离散余弦变换，对该帧进行重构，然后对重构后的宏块经过 JVT 处理模块 4 的环滤波模块 JC 进行环滤波，送入帧缓存区。重构的帧将用作以后预测编码帧的参考帧。反之，若宏块的编码模式是非帧内编码的，则开关 S 接 S2，首先对该宏块经过 JVT 处理模块 4 的运动估计模块 SD 进行运动估计，经过 JVT 处理模块 4 的运动补偿 SA，然后经 JVT 处理模块 4 的离散余弦变换/量化模块 JL 进行编码，同时进行逆量化/反离散余弦变换，在加法器 a2 进行重构；然后，经过 JVT 处理模块 4 的环滤波模块 JC 进行环滤波送入帧缓冲区 JB。经过帧内预测或经过 JVT 处理模块 4 的运动补偿 SA 后则可以得到当前宏块的活动性，送给率失真最优化模式选择模块 1，用以确定当前宏块的量化参数，然后，JVT 处理模块 4 根据码率控制的量化参数重新进行模式选择及编码。

最后应说明的是：以上实施例仅用以说明本发明而非限制，尽管参照较佳实施例对本发明进行了详细说明，本领域的普通技术人员应当理解，可以对本发明进行修改或者等同替换，而不脱离本发明的精神和范围，其均应涵盖在本发明的权利要求范围当中。

权利要求书

1、一种结合率失真优化的码率控制的方法，其特征在于：至少包括如下的步骤：

步骤 10：为 GOP 中各帧分配相应的比特数；

步骤 20：针对帧内各宏块进行率失真最优化模式选择，获得码率控制的量化参数；并将该量化参数输出，用于码率控制。

2、根据权利要求 1 所述的结合率失真优化的码率控制的方法，其特征在于：步骤 10 的具体操作为：在对 GOP 中各帧进行编码之前，根据平均帧的大小为 GOP 中各帧分配相应的比特数。

3、根据权利要求 2 所述的结合率失真优化的码率控制的方法，其特征在于：所述的平均帧的大小根据如下公式计算：

$R/F = R \div F$ ，其中：

R 为码率控制的目标码率；

F 为编码帧率；

R/F 为平均帧的大小的值。

4、根据权利要求 1 或 2 所述的结合率失真优化的码率控制的方法，其特征在于：所述的为 GOP 中各帧分配相应的比特数还进一步包括：在该 GOP 中的各帧之间进行协调，具体的协调处理依据如下的公式：

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_p X_b}{K_b X_p}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_b X_p}{K_p X_b}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

其中，

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数；

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目；

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值, 其为编码一帧所用的比特数与平均量化参数的乘积;

bit_rate 为目标码率;

$picture_rate$ 为编码帧率;

K_p 和 K_b 为经验值, 分别为 P, B 帧与 I 帧间的复杂性比例, 其取值范围为不小于 1.0 且不大于 2.0;

R 为当前 GOP 剩余的比特数; 且在编码一帧后, 依据如下公式对 R 进行更新:

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

5、根据权利要求 4 所述的结合率失真优化的码率控制的方法, 其特征在于: 在对一个 GOP 进行编码之前, 还依据如下公式对为当前 GOP 剩余的比特数进行初始化;

$$R = G + R_{prev}$$

$$G = bit_rate \times N \div picture_rate$$

其中:

R 为当前 GOP 剩余的比特数,

N 为当前 GOP 所包含的帧数,

G 为码率= bit_rate 时, 一个包含 N 帧图像的 GOP 的比特数,

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数, 且对于第一个 GOP,

$$R_{prev} = 0;$$

bit_rate 为目标码率;

$picture_rate$ 为编码帧率。

6、根据权利要求 4 所述的结合率失真优化的码率控制的方法, 其特征在于: X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定:

$$X_i = a \times bit_rate$$

$$X_p = b \times bit_rate$$

$$X_b = c \times bit_rate$$

其中:

a 、 b 、 c 为经验值, 其取值范围分别为: 100/115-160/115、20/115-60/115、10/115-

42/115;

bit_rate 为目标码率。

7、根据权利要求 1 所述的结合率失真优化的码率控制的方法，其特征在于：所述的步骤 20 至少包括第一次率失真最优化模式选择，即：用当前宏块之前一宏块的量化参数作为预测值进行模式选择；具体为：依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式：

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中：

s 为源图像当前宏块的辉度值；

c 为重构图像的当前宏块的辉度值；

λ_{MODE} 为拉格朗日常数；

对于 I 帧或 P 帧， $\lambda_{MODE} = 0.85 \times 2^{Q_{m-1}/3}$ ；

对于 B 帧， $\lambda_{MODE} = 4 \times 0.85 \times 2^{Q_{m-1}/3}$ ；

$D(s, c, MODE | QP)$ 为衡量当前宏块压缩后的失真率；

$R(s, c, MODE | QP)$ 为编码当前宏块所用得比特数；

QP 为当前宏块的量化参数。

8、根据权利要求 7 所述的结合率失真优化的码率控制的方法，其特征在于：当对 P 帧或 B 帧进行估计时，先进行运动估计，并选择使下式最小的运动向量为当前宏块的运动向量：

$$J(m, \lambda_{MOTION}) = SA(T) D(s, c(m)) + \lambda_{MOTION} R(m - p)$$

其中，

$D(s, c(m))$ 为衡量运动补偿的失真率的函数；

$SA(T) D$ 为运动补偿（或并经过哈德曼变换）后的宏块系数的绝对值和；

$R(m-p)$ 为编码运动向量所用的比特数；

s 为源图像的当前宏块的辉度值；

c 为预测图像的辉度值；

m 为运动向量；

p 为运动向量的预测值；

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

9、根据权利要求 8 所述的结合率失真优化的码率控制的方法, 其特征在于: 在第一次率失真最优化模式选择之后还进一步包括: 步骤 30: 当前宏块量化参数的确定; 具体为: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

10、根据权利要求 9 所述的结合率失真优化的码率控制的方法, 其特征在于: 根据当前宏块的活动性调整宏块的量化参数的处理具体为: 采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或运动补偿后的绝对值和作为宏块的活动性标志, 当前宏块的空间活动性依据下式计算:

$$N_act_m = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中,

i 为像素的横坐标, 并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标;

j 为像素的纵坐标, 并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标;

N_act_m 为当前宏块的空间活动性;

$s(i, j)$ 为源像素 (i, j) 的辉度值;

$c(i, j)$ 为像素 (i, j) 的预测值;

avg_act 为先前编码的同类型帧的平均 act_m 活动性;

act_m 为经过运动补偿或帧内预测后的宏块 m 的绝对值和。

11、根据权利要求 9 所述的结合率失真优化的码率控制的方法, 其特征在于: 所述的建立虚拟缓冲区模型具体为: 建立虚拟缓冲区的饱和度与宏块量化参数的映射关系; 其中, 当前宏块的量化参数由如下公式确定:

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_act_m$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times \text{bit_rate} / \text{picture_rate}$$

其中,

Q_n 为当前宏块的量化参数值;

d_n^n 为编码当前宏块时当前虚拟缓冲区的饱和度; 其中, 对于 I, P, B 帧, 其取值分别为 d_n^i 、 d_n^p 、 d_n^b ;

B_{n-1} 为编码前一宏块所用的比特数;

D_0^n 为编码该帧时虚拟缓冲区的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b ;

r 为虚拟缓冲区的大小值;

虚拟缓冲区最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m = MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

12、根据权利要求 11 所述的结合率失真优化的码率控制的方法, 其特征在于: 所述的虚拟缓冲区饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区的初始饱和度;

K_p 为 I, P 帧复杂性比例常数;

K_b 为 I, B 帧复杂性比例常数。

13、根据权利要求 9 或 10 或 11 或 12 所述的结合率失真优化的码率控制的方法, 其特征在于: 步骤 40: 在确定当前宏块的量化参数后, 还进一步包括第二次率失真最优化模式选择, 即: 用确定后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择; 具体依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像的当前宏块的辉度值, c 为重构图像的当前宏块的辉度值, λ_{MODE} 为拉格朗日常数;

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{Q_n - 1/3}$;

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{Q_{m-1}/3}$;

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

14、根据权利要求 13 所述的结合率失真优化的码率控制的方法, 其特征在于: 当对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m-p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数;

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和;

$R(m-p)$ 为编码运动向量所用的比特数;

s 为源图像的当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

15、一种结合率失真优化的码率控制的装置, 其特征在于: 它至少包括 JVT 处理模块、率失真最优化模式选择模块、虚拟缓冲区模块和全局复杂性估计模块; 其中,

JVT 处理模块输入帧数据, 并与率失真最优化模式选择模块、虚拟缓冲区模块和全局复杂性估计模块连接;

该率失真最优化模式选择模块分别与虚拟缓冲区和全局复杂性估计模块连接, 它从 JVT 处理模块输入信号, 并根据虚拟缓冲区模块和全局复杂性估计模块输出的信号将其输出信号回送给 JVT 处理模块, 并经过 JVT 处理模块处理后输出相应的量化参数信号。

16、根据权利要求 15 所述的结合率失真优化的码率控制的装置, 其特征在于: 所述的位分配为: 帧数据输入时, 首先为 GOP 中各帧分配相应的的比特数; 具体是: 在对 GOP 中各帧进行编码之前, 根据平均帧的大小为 GOP 中各帧分配相应的的比特数。

17、根据权利要求 16 所述的结合率失真优化的码率控制的装置，其特征在于：所述的平均帧的大小根据如下公式计算：

$R/F=R \div F$ ，其中：

R 为码率控制的目标码率；

F 为编码帧率；

R/F 为平均帧的大小的值。

18、根据权利要求 16 或 17 所述的结合率失真优化的码率控制的装置，其特征在于：所述的为 GOP 中各帧分配相应的比特数还包括：在该 GOP 中的各帧之间进行协调，具体的协调处理依据如下的公式：

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p X_p}{K_p X_i} + \frac{N_b X_b}{K_b X_i}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p K_p X_p}{K_p X_b}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b K_b X_b}{K_b X_p}}, \frac{bit_rate}{8 \times picture_rate} \right\}$$

其中，

T_i 、 T_p 和 T_b 分别为分配给 I 帧、P 帧和 B 帧的比特数；

N_i 、 N_p 和 N_b 分别为编码当前帧时 GOP 中包括的当前帧的剩余未编码的 I 帧、P 帧和 B 帧的数目；

X_i 、 X_p 和 X_b 分别为 I 帧、P 帧和 B 帧的全局复杂性估计值，其为编码一帧所用的比特数与平均量化参数的乘积；

bit_rate 为目标码率；

$picture_rate$ 为编码帧率；

K_p 和 K_b 为经验值，分别为 P、B 帧与 I 帧间的复杂性比例，其取值范围为不小于 1.0 且不大于 2.0；

R 为当前 GOP 剩余的比特数；且在编码一帧后，依据如下公式对 R 进行更新：

$$R = R - S_{i,p,b}$$

$S_{i,p,b}$ 为编码当前帧所用的比特数。

19、根据权利要求 18 所述的结合率失真优化的码率控制的装置，其特征在于：在对一个 GOP 进行编码之前，还依据如下公式对当前 GOP 剩余的比特数进行初始化；

$$R = G + R_{prev}$$

$$G = bit_rate \times N \div picture_rate$$

其中：

R 为当前 GOP 剩余的比特数；

N 为当前 GOP 所包含的帧数；

G 为码率= bit_rate 时，一个包含 N 帧图像的 GOP 的比特数；

R_{prev} 为编码前一 GOP 所剩余的或超出预分配的比特数，且对于第一个 GOP，

$$R_{prev} = 0;$$

bit_rate 为目标码率；

$picture_rate$ 为编码帧率。

20、根据权利要求 18 所述的结合率失真优化的码率控制的装置，其特征在于： X_i 、 X_p 和 X_b 的初始值分别依照如下公式确定：

$$X_i = a \times bit_rate$$

$$X_p = b \times bit_rate$$

$$X_b = c \times bit_rate$$

其中：

a 、 b 、 c 为经验值，其取值范围分别为：100/115-160/115、20/115-60/115、10/115-42/115；

bit_rate 为目标码率。

21、根据权利要求 15 所述的结合率失真优化的码率控制的装置，其特征在于：率失真最优化模式选择模块采用当前宏块之前一宏块的量化参数作为预测值进行模式选择；其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式：

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中：

s 为源图像的当前宏块的辉度值;

c 为重构图像的当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数;

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{\frac{Q_{m-1}}{3}}$;

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{\frac{Q_{m-1}}{3}}$;

$D(s, c, MODE/QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE/QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

22、根据权利要求 21 所述的结合率失真优化的码率控制的装置, 其特征在于: 率失真最优化模式选择模块对 P 帧或 B 帧进行估计时, 先进行运动估计, 并选择使下式最小的运动向量为当前宏块的运动向量:

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m-p)$$

其中,

$D(s, c(m))$ 为衡量运动补偿的失真率的函数;

$SA(T)D$ 为运动补偿 (或并经过哈德曼变换) 后的宏块系数的绝对值和;

$R(m-p)$ 为编码运动向量所用的比特数;

s 为源图像的当前宏块的辉度值;

c 为预测图像的辉度值;

m 为运动向量;

p 为运动向量的预测值;

λ_{MOTION} 为拉格朗日常数, 并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$; 其中,

λ_{MODE} 为拉格朗日常数。

23、根据权利要求 22 所述的结合率失真优化的码率控制的装置, 其特征在于: 在率失真最优化模式选择模块第一次率失真最优化模式选择之后, 其输出给 JVT 处理模块的信号还由 JVT 处理模块进一步当前宏块的量化参数进行确定; 具体为: 根据当前宏块的活动性调整宏块的量化参数并建立虚拟缓冲区模型。

24、根据权利要求 22 所述的结合率失真优化的码率控制的装置, 其特征在于: 上述

的根据当前宏块的活动性调整宏块的量化参数的处理具体为：率失真最优化模式选择模块采用经过第一次率失真最优化模式选择之后得到的、当前宏块经过帧内预测或运动补偿后的绝对值和作为宏块的活动性标志，

当前宏块的空间活动性依据下式计算：

$$N_{-act_m} = \frac{(2 \times act_j) + avg_act}{act_j + (2 \times avg_act)} \quad act_m = \sum_{i,j} |s(i,j) - c(i,j)|$$

其中，

i 为像素的横坐标，并且其取值范围为当前宏块的左上角顶点的横坐标至当前宏块的右下角顶点的横坐标；

j 为像素的纵坐标，并且其取值范围为当前宏块的左上角顶点的纵坐标至当前宏块的右下角顶点的纵坐标；

N_{-act_m} 为当前宏块的空间活动性；

$s(i, j)$ 为源像素 (i, j) 的辉度值；

$c(i, j)$ 为像素 (i, j) 的预测值；

avg_act 为先前编码的同类型帧的平均 act_m 活动性；

act_m 为经过运动补偿或帧内预测后的当前宏块的绝对值和。

25、根据权利要求 22 所述的结合率失真优化的码率控制的装置，其特征在于：所述的建立虚拟缓冲区模型具体为：虚拟缓冲区模块建立虚拟缓冲区的饱和度与宏块量化参数的映射关系；其中，当前宏块的量化参数由如下公式确定：

$$Q_m = \left(\frac{d_m^n \times 31}{r} \right) \times N_{-act_m}$$

$$d_m^n = d_0^n + B_{m-1} - T_n \times (m-1) / MB_CNT$$

$$r = 2 \times bit_rate / picture_rate$$

其中，

Q_m 为当前宏块的量化参数值；

d_m^n 为编码当前宏块时当前虚拟缓冲区的饱和度；其中，对于 I, P, B 帧，其取值分别为 d_m^I 、 d_m^P 、 d_m^B ；

B_{n-1} 为编码前一宏块所用的比特数;

D_0^n 为编码该帧时虚拟缓冲区的初始饱和度, n 的值为 i 或 p 或 b , 对应的初始饱和度为 d_0^i 或 d_0^p 或 d_0^b ; 且在编码下一帧时为新的初始饱和度 d_0^i 或 d_0^p 或 d_0^b ;

r 为虚拟缓冲区的大小值;

虚拟缓冲区最终的饱和度为 d_m^i 或 d_m^p 或 d_m^b ; 其中: $m=MB_CNT$;

MB_CNT 为一帧图像所包含的宏块数。

26、根据权利要求 25 所述的结合率失真优化的码率控制的装置, 其特征在于: 所述的虚拟缓冲区饱和度的初始值在编码第一帧时为:

$$d_0^i = 10 \times r / 31 \quad d_0^p = K_p \times d_0^i \quad d_0^b = K_b \times d_0^i$$

其中,

r 为虚拟缓冲区的大小值;

d_0^i 或 d_0^p 或 d_0^b 分别为 i 帧或 p 帧或 b 帧虚拟缓冲区的初始饱和度;

K_p 为 I , P 帧复杂性比例常数;

K_b 为 I , B 帧复杂性比例常数。

27、根据权利要求 23 或 24 或 25 或 26 所述的结合率失真优化的码率控制的装置, 其特征在于: 在确定当前宏块的量化参数后, JVT 处理模块将该参数信号输出给率失真最优化模式选择模块, 由该率失真最优化模式选择模块进行率失真最优化模式选择, 即: 用确定后的当前宏块量化参数对当前宏块再次进行率失真最优化模式选择; 其依照如下公式计算并选择其最小的编码模式为当前宏块的初始编码模式:

$$D(s, c, MODE | QP) + \lambda_{MODE} R(s, c, MODE | QP)$$

其中:

s 为源图像当前宏块的辉度值;

c 为重构图像当前宏块的辉度值;

λ_{MODE} 为拉格朗日常数;

对于 I 帧或 P 帧, $\lambda_{MODE} = 0.85 \times 2^{Q_{m-1}/3}$;

对于 B 帧, $\lambda_{MODE} = 4 \times 0.85 \times 2^{Q_{m-1}/3}$;

$D(s, c, MODE | QP)$ 为衡量当前宏块压缩后的失真率;

$R(s, c, MODE | QP)$ 为编码当前宏块所用得比特数;

QP 为当前宏块的量化参数。

28、根据权利要求 27 所述的结合率失真优化的码率控制的装置，其特征在于：当率失真最优化模式选择模块对 P 帧或 B 帧进行估计时，先进行运动估计，并选择使下式最小的运动向量为当前宏块的运动向量：

$$J(m, \lambda_{MOTION}) = SA(T)D(s, c(m)) + \lambda_{MOTION} R(m - p)$$

其中；

$D(s, c(m))$ 为衡量运动补偿的失真率的函数；

$SA(T)D$ 为运动补偿（或并经过哈德曼变换）后的宏块系数的绝对值和；

$R(m-p)$ 为编码运动向量所用的比特数；

s 为源图像当前宏块的辉度值；

c 为预测图像的辉度值；

m 为运动向量；

p 为运动向量的预测值；

λ_{MOTION} 为拉格朗日常数，并且 $\lambda_{MOTION} = \sqrt{\lambda_{MODE}}$ ；其中，

λ_{MODE} 为拉格朗日常数。

29、根据权利要求 28 所述的结合率失真优化的码率控制的装置，其特征在于：率失真最优化模式选择模块根据虚拟缓冲区和全局复杂性估计模块输出的信号，将处理后的量化参数信号回送给 JVT 处理模块，并经过 JVT 处理模块处理后输出。

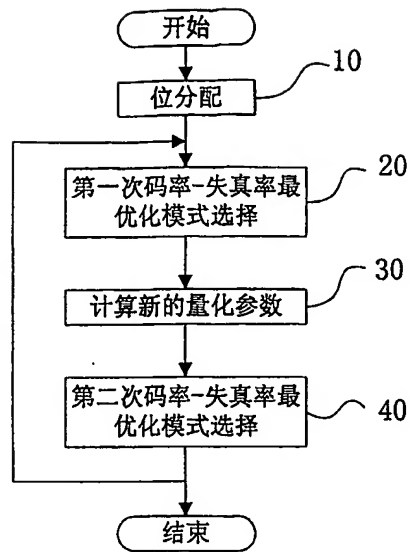


图 1

